视频与视觉技术国家工程研究中心
NATIONAL ENGINEERING RESEARCH CENTER OF VISUAL TECHNOLOGY

Faster and Stronger: When ANN-SNN Conversion Meets Parallel Spiking Calculation

Zecheng Hao, Qichao Ma, Kang Chen, Yi Zhang, Zhaofei Yu, Tiejun Huang

ICML 2025

北京大学
PEKING UNIVERSITY

Faster and Stronger: When ANN-SNN Conversion Meets Parallel Spiking Calculation

Zecheng Hao, Qichao Ma, Kang Chen, Yi Zhang, Zhaofei Yu, Tiejun Huang

ICML 2025

Github Code

Abstract

Spiking Neural Network (SNN), as a brain inspired and energy-efficient network, is currently facing the pivotal challenge of exploring a suitable and efficient learning framework. In this work, we propose a novel parallel conversion learning framework, which establishes a mathematical mapping relationship between each time-step of the parallel spiking neurons and the cumulative spike firing rate. Furthermore, by integrating the above framework with the distribution-aware error calibration technique, we can achieve efficient conversion towards more general activation functions or training-free circumstance.

Theorems

$$\mathbf{\Lambda}^l_{\text{POST}} = \begin{bmatrix} \mathrm{c}^{l,1} \\ \mathrm{c}^{l,2} \\ \vdots \\ \mathrm{c}^{l,T} \end{bmatrix} \odot \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}.$$

$$\begin{cases} x \cdot \mathrm{c}^{l,x} \cdot \mathbf{W}^l \mathbf{r}^{(l-1),T} = \theta^l \\ \mathbf{W}^l \mathbf{r}^{(l-1),T} T = (T-x+1)\theta^l \end{cases} \Rightarrow \mathrm{c}^{l,x} = \frac{T}{x(T-x+1)}.$$

$$\mathbf{\Lambda}^l_{\text{PC}} = \mathbf{\Lambda}^l_{\text{POST}} \mathbf{\Lambda}^l_{\text{PRE}} = \begin{bmatrix} \frac{1}{T} & \frac{1}{T} & \cdots & \frac{1}{T} \\ \frac{1}{T-1} & \frac{1}{T-1} & \cdots & \frac{1}{T-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}.$$

Methodology

Table 1: Comparison of various supervised learning methods for SNNs.

Method	Act. Func.	Train. Free	Train. Speed	Train. Mem.	Inf. Lat.	Inf. Speed	Inf. Acc.
STBP Training	Surro. Func.	✗	Slow	Large	Ultra Low	Slow	Low
ANN-SNN Conversion	ReLU	✗	Fast	Small	Ultra High	Slow	High
	QCFS	✗	Fast	Small	High	Slow	High
Conversion Rect.	QCFS	✗	Fast	Small	Low	Slow	High
This Work	QCFS	✗	Fast	Small	Ultra Low	Fast	High
	ReLU	✓	N/A	N/A	Low	Fast	High

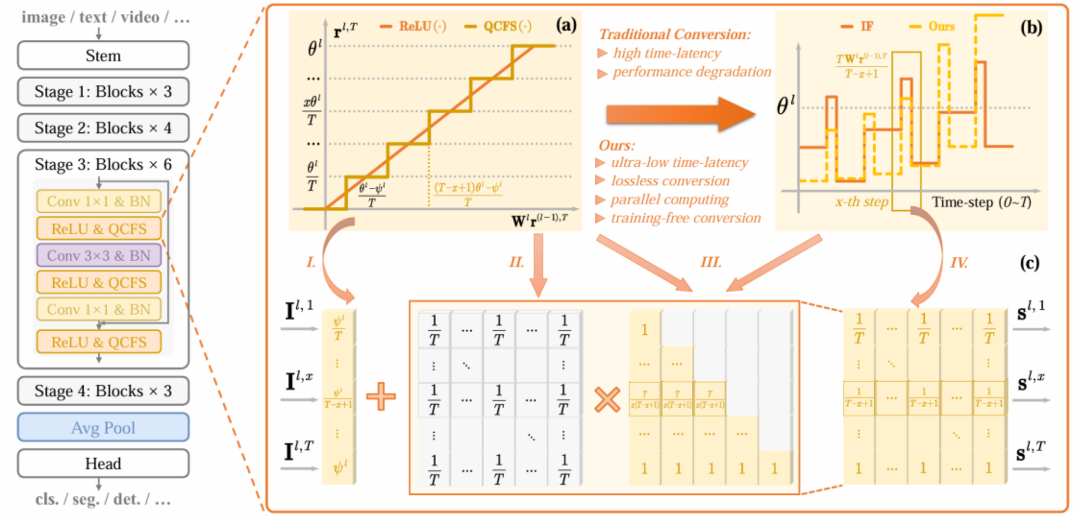


Figure 1: The overall framework of parallel conversion. Here (a) depicts the activation functions in ANNs, (b) shows the sorting property of parallel spiking neurons in the firing phase, and (c) describes the specific process of parallel inference.

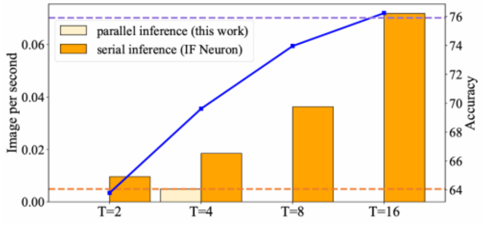
Table 2: Detailed experimental configuration for universal parallel conversion framework.

Conversion Cases	Need Thre. Rec.	Need Calib.	$\mathbf{\Lambda}^l_{\text{PC}}$.shape	$\mathbf{b}^l$.shape	θ^l_{PRE}.shape	θ^l_{POST}.shape
QCFS ($\tilde{T} = T$)	✗	✗	$[T, T]$	$[T,]$	scalar	scalar
QCFS ($\tilde{T} \neq T$)	✗	✓	$[T, T]$	$[T, C]$	scalar	$[C,]$
ReLU	✓	✓	$[T, T]$	$[T, C]$	$[C,]$	$[C,]$

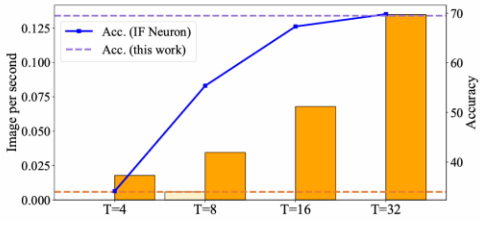
Experiments

Table 3: Comparison of previous state-of-the-art learning methods. † denotes adopting the error calibration technique.

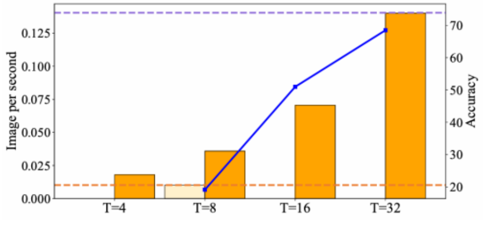
Dataset	Method	Type	ANN Acc.(%)	Arch.	T	SNN Acc.(%)
CIFAR-10	OPT	ANN-SNN Conversion	93.51	VGG-16	32	88.79
	QCFS	ANN-SNN Conversion	95.52	VGG-16	2, 4, 8	91.18, 93.96, 94.95
	SNM	Conversion Rect.	94.09	VGG-16	32	93.43
	SRP	Conversion Rect.	95.52	VGG-16	6 (4+2)	94.47
	Ours	**Parallel Conversion**	95.43	VGG-16	2, 4	**94.16, 95.50**
	QCFS	ANN-SNN Conversion	91.77	ResNet-20	2, 4, 8	73.20, 83.75, 89.55
	SRP	Conversion Rect.	91.77	ResNet-20	6 (4+2)	88.73
	Ours	**Parallel Conversion**	91.67	ResNet-20	2, 4	**87.42, 91.58**
CIFAR-100	OPT	ANN-SNN Conversion	70.21	VGG-16	32	56.16
	QCFS	ANN-SNN Conversion	76.28	VGG-16	2, 4, 8	63.79, 69.62, 73.96
	SNM	Conversion Rect.	74.13	VGG-16	32	71.80
	SRP	Conversion Rect.	76.28	VGG-16	6 (4+2)	74.31
	Ours	**Parallel Conversion**	76.11	VGG-16	2, 4	**72.71, 75.98**
	QCFS	ANN-SNN Conversion	69.94	ResNet-20	4, 8, 16	34.14, 55.37, 67.33
	SRP	Conversion Rect.	69.94	ResNet-20	6 (4+2)	53.96
	Ours	**Parallel Conversion**	69.57	ResNet-20	4, 8	**65.31, 69.62**
ImageNet-1k	OPT	ANN-SNN Conversion	72.40	VGG-16	32	54.92
	QCFS	ANN-SNN Conversion	74.29	VGG-16	8, 16, 32	19.12, 50.97, 68.47
	SNM	Conversion Rect.	73.18	VGG-16	32	64.78
	Burst	Conversion Rect.	74.27	VGG-16	32	70.61
	COS	Conversion Rect.	74.19	VGG-16	10 (8+2)	70.59
	Ours	**Parallel Conversion**	74.23	VGG-16	4, 8, 16	**71.23, 73.92, 74.26**
	Ours†	**Parallel Conversion**	74.23	VGG-16	4	**71.75**
	RecDis	STBP Training	-	ResNet-34	6	67.33
	Dspike	STBP Training	-	ResNet-34	6	68.19
	GLIF	STBP Training	-	ResNet-34	4	67.52
	TAB	STBP Training	-	ResNet-34	4	67.78
	OPT	ANN-SNN Conversion	70.95	ResNet-34	64	59.52
	QCFS	ANN-SNN Conversion	74.32	ResNet-34	8, 16, 32	35.06, 59.35, 69.37
	COS	Conversion Rect.	74.22	ResNet-34	10 (8+2)	72.66
	Ours	**Parallel Conversion**	74.30	ResNet-34	4, 8	**67.28, 74.32**
	Ours†	**Parallel Conversion**	74.30	ResNet-34	4	**72.90**



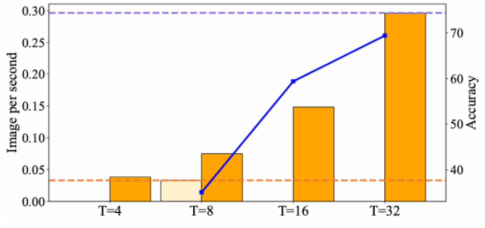
(a) CIFAR-100, VGG-16



(b) CIFAR-100, ResNet-20



(c) ImageNet-1k, VGG-16



(d) ImageNet-1k, ResNet-34

PEKING UNIVERSITY 1898

Thanks for Listening!